SECURITY CLASSIFICATION OF THIS PAGE

AD-A243 716

REPORT DOCUMENTATION PAGE

1b. RESTRICTIVE MARKINGS	
3. DISTRIBUTION/AVAILABILITY OF REPORT	Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR-
6a. NAME OF PERFORMING ORGANIZATION: Department of Psychology, Yale University	6b. OFFICE SYMBOL (If applicable)
7a. NAME OF MONITORING ORGANIZATION	Air Force Office of Scientific Research
6c. ADDRESS (City, State and ZIP Code)	New Haven, CT 06520
7b. ADDRESS (City, State and ZIP Code)	Bolling Air Force Base, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR	8b. OFFICE SYMBOL (If applicable): NL
9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER	AFOSR-91-0020
8c. ADDRESS (City, State and ZIP Code)	Building 410, Bolling AFB DC 20332-6448

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2313	A6	

11. TITLE (Include Security Classification) Signal- and Listener-based Factors in Complex Auditory Pattern Perception

12. PERSONAL AUTHOR(S) Arthur G. Samuel

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
FINAL	FROM 10/1/90 TO 9/30/9[illegible]	1991-Oct. -7	21

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The research conducted during the one year funding period was a subset of the original three-year study of the perception of complex auditory patterns, including speech and music. One set of experiments explored two early stages in the perception of complex signals, using adaptation procedures. This research investigated effects of varying signal amplitude, and the effects of more cognitive factors: lexical knowledge, and the listener's level of attention to the adapting sound. A second set of experiments examined perceptual restoration effects. Those experiments investigated how knowledge of particular words influenced the perceptual restoration of deleted or degraded portions of the word. The two lines of research represent progress toward understanding the analyses conducted on complex auditory patterns by human listeners.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED XX SAME AS RPT XX DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: Dr. John Tangney	22b. TELEPHONE NUMBER (Include Area Code): (202) 767-5021
22c. OFFICE SYMBOL	NL

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE. UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

1 8 OCT

I. Summary

The research conducted during the one year funding period was a subset of the original three-year study of the perception of complex auditory patterns, including speech and music. One set of experiments explored two early stages in the perception of complex signals, using adaptation procedures. This research investigated effects of varying signal amplitude, and the effects of more cognitive factors: lexical knowledge, and the listener's level of attention to the adapting sound. A second set of experiments examined perceptual restoration effects. Those experiments investigated how knowledge of particular words influenced the perceptual restoration of deleted or degraded portions of the word. The two lines of research represent progress toward understanding the analyses conducted on complex auditory patterns by human listeners.

II. Research Objectives

The objective of the research project is to delineate principles that underlie the perception of complex auditory patterns. The stimuli used are speech and musical patterns of varying complexity. A wide array of experimental procedures and analyses are used to try to determine properties that are true of the perception of complex auditory patterns across stimulus domains. In addition, we also are interested in discovering any principles that are domain specific (e.g., as "categorical perception" has traditionally been claimed to be a principle of perception specific to the speech domain). The various experimental investigations in the project may be broadly grouped into studies of signal-based factors, and studies of listener-based factors. The former group includes experiments that explore how properties of the input signal determine perception, while the latter group includes studies of how listeners' expectations influence perception/performance. The former group primarily focusses on early representations of the signal, and the latter includes higher-level factors (including, but not limited to, attentional influences). The long-term goal of the research is to understand both signal-based and listener-based factors, and their interaction in the perception of complex auditory patterns.

III. Research Completed During the Funding Period

During the one year period covered in this report, we have moved our laboratory from Yale University to the State University of New York at Stony Brook. This transfer of the laboratory effectively shortened the period of productive time by several months. Nevertheless, we have conducted research along four lines, including two from the original proposal and two that have developed since then. These projects are ongoing, in various stages of completion. In this Section, we will briefly outline the progress made on each project during the funding period. Work that is at or near completion is listed in Section IV (Publications), and in Section VI (Interactions/Conference Presentations).

A. Early representational levels. During the funding period, we have continued our research program that focuses on the early codings of complex acoustic patterns. This research line uses the selective adaptation paradigm, in which repeated presentation of a sound (the "adaptor") induces changes in a listener's identification of related sounds. Our previous work in this research program has produced both empirical and theoretical gains. Empirically, we have specified important acoustic properties (and combinations of properties) in the perception of complex auditory patterns, including both speech and nonspeech sounds (e.g., Kat and Samuel, 1984; Samuel, 1988; Samuel and Newport, 1979). Theoretically, we developed a model of the early processing of such sounds which included two qualitatively different levels of analysis (e.g., Samuel, 1986, 1988; see also Samuel 1977), and which characterizes the nature of units at each level (e.g., Samuel, 1982, 1989). In the two-level model, the first level is hypothesized to be rather like a neural spectrogram, with degree of adaptation being a function of spectral overlap; it is relatively peripheral (primarily monaurally-driven), subject to neural fatigue, and non-domain-specific (e.g., not specific to speech). The second level involves more abstract or encoded representations for which spectral overlap is less important in determining adaptation effects; these units are more central (predominantly binaurally-driven), primarily subject to criterion shifts rather than fatigue, and like the first-level units, non-domain-specific. Thus these units can provide phonetic codings, but are not inherently linguistic; non-speech signals can engage these "phonetic" units and alter their functioning (e.g., Samuel, 1988; Samuel and Newport, 1979).

In a recent series of experiments (Kat and Samuel, 1988) we tested the hypothesis that adaptation-induced fatigue produces slower responding to stimuli in the adapted category, while adaptation-induced criterion shifts do not. The test involved having subjects rapidly label stimuli from an 8-step /ba-da/ continuum, under four critical adaptation conditions. Two adaptors were tokens that included the same second and third formants as the endpoint /ba/ or /da/, but without F1. These adaptors contained much of the critical acoustic information for /ba/ or for /da/, but had very weak phonetic information. In our model, they should engage first-level units for /ba/ or /da/ rather well because of the good spectral overlap, but have little effect on second-level categorization because of the missing phonetic quality. These "F2F3" adaptors were compared to the adaptors /pa/ and /ta/, which share much phonetically with /ba/ and /da/, respectively, but which are acoustically rather different because of the differences in voicing. The /pa/ and /ta/ adaptors could be expected to engage second-level units for /ba/ and /da/ rather well, but not to affect relevant first-level units very much. The overall labeling of /ba-da/ syllables was equally affected by the two types of adaptors (F2F3 versus /pa/ and /ta/). However, only the spectrally-matched F2F3 adaptors caused reliable reaction time changes. Note that this is exactly as predicted by the model, since the F2F3 stimuli are hypothesized to fatigue first-level units, and thereby slow responses, while the /pa/ and /ta/ adaptors are hypothesized to induce (non-slowing) criterion shifts.

However, one aspect of the data contradicted the view that the F2F3 adaptor was affecting units at an early, peripheral, level: The adaptation effects were not reduced when the adaptor was presented contralaterally to the test items; the affected units are binaurally driven. This has led us to postulate an intermediate level in the model, a binaurally driven level that is subject to fatigue. These new findings made one of our planned experiments particularly interesting -- an experiment varying the amplitude of our F2F3 and /pa/-/ta/ adaptors. In our grant proposal, we had predicted that first-level units should be relatively amplitude-sensitive, while second-level units should not: If, as the model and data suggest, first-level units are subject to neural fatigue, then louder adaptors should induce more fatigue than softer ones. In contrast, criterion-shift effects should not vary with adaptor amplitude. Thus, the test involved using the F2F3 and /pa/ and /ta/ adaptors on the /ba-da/ test series, using different adaptor amplitudes.

We have conducted this test recently, and the results were counter to the original prediction, and instead consistent with our developing understanding of the intermediate level of representation. The F2F3 adaptors, which we are hypothesizing engage this middle level, did not produce amplitude-dependent effects. This result is consistent with our new conception of this level as a central, binaurally-driven site. The interesting issue to resolve at this point is how to characterize a level of representation that appears to be central, but shows apparent fatigue-related properties. In particular, it will be critical to specify this level's relationship to our hypothesized earlier level (monaurally-driven, subject to fatigue) and hypothesized next level (binaurally-driven, subject to criterion shifts).

We have conducted two other adaptation studies in this period, both of which were designed to look for upper bounds to the effects. More specifically, we were testing for cognitive influences on these lower level representations. One study examined possible lexical influences, and the other tested whether the adaptation effects require active attention to the adaptor.

Two versions of the lexical test have been run. In one experiment, we tested the hypothesis that words are perceptually more stable than comparable nonwords. On this view, a word (e.g., "gift" or "kiss") should resist adaptation more than a similar nonword (e.g., "giss" or "kift"). To test this hypothesis, we conducted adaptation experiments with "gi" and "ki", and measured identification of syllables from a "gift-kift" and a "giss-kiss" continuum. We found no evidence for lexical codes resisting adaptation -- identification shifts were just as large when "gift" and "kiss" report was reduced as when "giss" and "kift" were. These data indicate that the levels tapped by the adaptation procedure are below any effects produced by lexical codes.

The second study of possible cognitive influences on adaptation produced results consistent with the lexical experiments. We compared adaptation effects under the usual testing conditions to a situation in which subjects were required to do a demanding second task during each adaptation period. In a typical adaptation experiment, adaptation phases of about a half minute alternate with test trials (when subjects identify the test syllables). In the critical experimental condition, subjects were given a running arithmetic problem to work on during each 30-second adaptation period. If processing the adaptor requires cognitive resources, then adaptation effects should be weaker in the dual-task condition than in the standard case. In fact, the distractor did not weaken the adaptation effects. This result, like the results for the lexical study, indicates that the structures being tapped in our adaptation experiments are indeed the low level ones that we have hypothesized -- they are immune to variations in cognitive load. We are continuing this line of research, with a goal of specifying the operation of units at the three early levels that we have identified.

B. Perceptual Restoration. The perceptual system is designed to produce a filtered version of reality: Incomplete or ambiguous stimuli will usually be perceived as more complete and less ambiguous than the input. The operation of such restoration processes makes it clear that a full understanding of perception must extend beyond the specification of signal-based factors. Studies of perceptual restoration effects, beginning with Warren's (1970) seminal paper, have begun to to clarify the perceptual architecture. The discrimination methodology introduced by Samuel (1981) has proven very useful in distinguishing between perceptual restoration and post-perceptual biases; these results have played an important role in the debate over modular versus interactive architectures (cf Fodor, 1983).

In the discrimination paradigm, two types of stimuli are constructed. In one type, a segment of the speech is digitally removed, and replaced by white noise of approximately the same amplitude. In the second stimulus type, the noise is superimposed on the corresponding segment. Listeners must judge whether a given stimulus is intact (i.e., has noise superimposed on a speech segment), or not (i.e., has noise replacing a speech segment). Of course, if the perceptual system restores deleted speech segments, then the two stimulus types will be difficult to discriminate because both will sound intact, with superimposed noise.

The discrimination methodology relies on signal detection analyses to separate perceptual effects (d') from postperceptual biases (Beta). However, an argument can be made for the view that true perceptual effects could show up in the bias parameter in some cases. Consider a manipulation such as stimulus lexicality -- the test item can be either a real word, or a matched pseudoword. Because <u>both</u> stimulus types (added and replaced) are degraded, it is at least plausible that both types could gain from the lexicality. In signal detection terms, both the "replaced" distribution and the "added" distribution would be shifted up the axis of "perceptual completeness". Note that such an effect would be mathematically equivalent to shifting the decision criteria (for reporting "intact") down the scale -- a Beta shift. We have conducted a study to test this hypothesis. The method involves presenting added and replaced stimuli, either words or pseudowords, followed by the intact, original word or pseudoword.

The final word or pseudoword is a standard that subjects used in judging the intactness of the added/replaced stimulus, using an eight-point rating scale of stimulus quality. If lexicality does in fact shift both distributions, then this should show up in the ratings: Both added and replaced stimuli should get higher intactness ratings relative to the fixed standard in the word condition than in the pseudoword case. Alternatively, the truly intact stimuli might not receive any increased sense of intactness via lexical information because such stimuli are already intact; in essence, they may be at ceiling on perceived intactness.

The results of our experiments using the new rating scale procedure have been very interesting. We have been able to use the technique to map out distributions of perceived intactness for the added and replaced versions of words and pseudowords. Using stimuli that have in earlier work produced discriminability differences between words and pseudowords, we have found corresponding differences in the distributions for words versus pseudowords: The intactness ratings for intact words do not differ greatly from those for intact pseudowords, while there is an appropriate difference in the distributions for stimuli with deleted segments. This pattern is exactly the one that the methodology was originally predicated on. In the next two months, we will be conducting follow-up experiments with the rating methodology, using a stimulus set that has failed to produce discriminability differences in the standard methodology. We expect that these stimuli will exhibit shifts in both the added and replaced distributions. We have conducted finer grained analyses, looking at factors such as phone class and temporal position of the critical phoneme, to determine when a given pattern of data will be found. The rating scale technique promises to provide a much more sensitive way to examine these critical issues.

C. Lexical influences on phoneme identification. We have developed two additional research lines that investigate lexical influences on the identification of speech sounds. One of these projects, done in collaboration with Mark Pitt (of Ohio State University), is a comprehensive set of experiments using Ganong's (1980) phoneme identification task. In Ganong's task, listeners classify a phoneme in a carrier word or nonword. For example, subjects would indicate whether the initial consonant was "d" or "t" when presented with items from a continuum ranging between "dash" and "tash", or from "dask" to "task". Note that in one continuum, "d" responses are consistent with a word ("dash"), while in the other, "t" responses are consistent with a word ("task"). A number of studies have reported identification and reaction time effects that seem to depend on such lexical consistency. However, there have been conflicting interpretations of these results, with some authors attributing the effects to top-down lexical influences, while others postulate non-interactive ("autonomous") models. We believe that much of the confusion stems from uncontrolled variation in testing conditions across studies. In our work, we have parametrically varied several of the factors that have varied across previous studies, including type of phonetic contrast (e.g., voicing versus place of articulation), stimulus quality (e.g., in the clear or in a noisy background), cognitive load (e.g., phoneme identification with or without a concurrent task), and position of the critical phoneme (e.g., utterance-initial versus utterance-final). We have found that these procedural factors must be considered in interpreting the results, especially the factor of phoneme position: Tests with utterance-initial critical phonemes produce results that are more consistent with non-interactive models, while tests using utterance-final phonemes are more easily accommodated by interactive models. This research project has been submitted for publication.

We are working on one conceptually similar research project, a project that developed as a result of the P.I.'s stint as a Visiting Scientist last year at the Max Planck Institute for Psycholinguistics in Nijmegen (the Netherlands). This work, done in collaboration with Uli Frauenfelder, uses stimuli similar to those in the research just discussed. The critical stimuli are tokens that are constructed to be ambiguously identifiable as, for example, either "dent" or "tent". In this research project, we are testing predictions of interactive models that involve inhibition of candidates by others. For example, in these models, the activation of one word is hypothesized to inhibit the activation of others. Thus, to the extent that "dent" is activated, it inhibits activation of "tent" (and vice versa). To test this prediction, we are constructing sets of stimuli in which there should be predictable variation in such inhibition. For example, identification of the "d" in "dent" should be slower than the "d" in "deaf", when the latter is constructed to be ambiguous between "deaf" and the nonword "tef". This follows from the fact that the nonword "tef" is not expected to inhibit "deaf" in the way that "tent" inhibits "dent". We have devised 40 sets of this sort, and are in the process of stimulus construction. When the stimuli are made and piloted, we will use them to contrast the predictions of interactive models (inhibition by competitors) to the predictions of autonomous models (independent activation of words). This research project fits well with the other projects we are conducting that aim to define the structure and operation of the system that deals with complex signals, including speech and music.

References

Fodor, J.A. (1983). The Modularity of Mind. MIT Press, Cambridge, Mass.

Ganong, W.F. (1980). Phonetic categorization in auditory word perception. Journal of Experimental Psychology: Human Perception and Performance, 6, 110-125.

Kat, D., and Samuel, A.G. (1984). More adaptation of speech by nonspeech. Journal of Experimental Psychology: Human Perception and Performance, 10, 512-525.

Kat, D., Samuel, A.G. (1988) Using reaction times to explore the mechanisms of selective adaptation. Journal of the Acoustical Society of America 84 (S), 160.

Samuel, A.G. (1977). The effect of discrimination training on speech perception: Noncategorical perception. Perception & Psychophysics, 22, 321-330.

Samuel, A.G. (1981). Phonemic restoration: Insights from a new methodology. Journal of Experimental Psychology: General, 110, 474-494.

Samuel, A.G. (1982). Phonetic prototypes. Perception &

Psychophysics, 31, 307-314.

Samuel, A.G. (1986). Red herring detectors and speech perception: In defense of selective adaptation. Cognitive Psychology, 18, 452-499.

Samuel, A.G. (1988). Central and peripheral representation of whispered and voiced speech. Journal of Experimental Psychology: Human Perception and Performance, 14, 379-388.

Samuel, A.G. (1989). Insights from a failure of selective adaptation: Syllable-initial and syllable-final consonants are different. Perception and Psychophysics, 45, 485-493.

Warren, R.M. (1970). Perceptual restoration of missing speech sounds. Science, 167, 392-393.

IV. Papers appearing during the funding period

Pitt, M.A., and Samuel, A.G. (1990). Attentional allocation during speech perception: How fine is the focus? Journal of Memory and Language, 29, 611-632.

Samuel, A.G. (1991). Perceptual degradation due to signal alternation: Implications for auditory pattern processing. Journal of Experimental Psychology: Human Perception and Performance, 17, 392-403.

Samuel, A.G. (1991). A further examination of the role of attention in the phonemic restoration illusion. Quarterly Journal of Experimental Psychology, 43A, 679-699.

V. Personnel

Principal Investigator: Arthur G. Samuel, Associate Professor of Psychology at Yale University. Ph.D. from University of California, San Diego, 1979.

Associate in Research: Donna Kat. B.A. in Psychology from University of California. San Diego, 1979.

VI. List of Interactions (coupling activities)

Conference Presentations

During the one-year funding period, we have made five presentations at three conferences. These are summarized here:

A. Samuel and D. Kat. (Fall 1990) Selective Adaptation Five Years Later: Bringing Home the Herring. Presented to the Psychonomic Society. New Orleans.

M. Pitt and A. Samuel. (Fall 1990) Is Phoneme Identification Facilitated by Feedback from a Words' Lexical Representation? Presented to the Psychonomic Society. New Orleans.

A Samuel. (Fall 1990) Signal alternation disrupts perception of music as well as speech. Presented to the Acoustical Society of America. San Diego.

D. Kat (Fall 1990) Reaction time and dichotic evidence for central processing of complex auditory signals. Presented to the Acoustical Society of America. San Diego.

A Samuel (Spring 1991) Perceptual degradation due to signal alternation. Presented at the SISSA Workshop on "The Psycholinguistic Consequences of Contrasting Language Phonologies", Trieste, Italy.

<u>Other Interactions</u>:

In addition to the conferences listed above, the PI has been heavily involved in the review process for both granting agencies and journals. This includes grant reviews for AFOSR and NSF, and manuscript reviews for:

Behavioral Research Methods and Instrumentation
Cognition
Cognitive Psychology
Journal of the Acoustical Society of America
Journal of Experimental Psychology: Human Perception and Performance
Language and Speech
Memory and Cognition
Perception & Psychphysics
Quaterly Journal of Experimental Psychology

During the granting period, the PI was a member of the Editorial Board of three journals: <u>Cognition</u>, <u>Memory and Cognition</u>, and the <u>Journal of Experimental Psychology</u>: <u>Human Perception and Performance</u>. These professional activities produce a great deal of interaction with other researchers. In addition, the P.I. has been continuing a collaborative project with Uli Frauenfelder of the Max Planck Institute (Nijmegen, the Netherlands).